

Ethics of “Taste, Ties, and Time”

Huanye Liu

The ethical issues encountered in scientific research are gaining more and more public attention these years, especially those concerning with scraping data from social networking websites, which has gradually become common practice when scientists want to gain access to large amount of first-hand data. The project conducted by the T3 research team is one such case.

Based on Salganik’s principles of ethical research, the project has violated each of the four principles to different extents. What follows is a brief analysis.

1 Respect for persons

Among the four principles, the T3 project falls short of the first principle most when the researchers released the data collected from the Facebook accounts of all members of the Class of 2009 at Harvard College. Although the research team claimed they had made attempts to protect the participants’ privacy, they ignored one important implication of the first principle of ethical research: receive informed consent from participants.

First of all, the T3 team employed several RAs from the source institution to collect relevant data because those RAs were in the same network as the participants. But the profile data of the participants were originally set with a privacy setting meant to share the data only within the network. In other words, the T3 team claimed that the data from Facebook were already public, but meanwhile they knew they had to employ RAs from within the network to collect the data they wanted. So the data collection was apparently against the participants’ intention to keep the data from leaking outside the network. Therefore, we can see that such behavior did not respect the participants’ expectation on their data privacy.

Secondly, neither the researchers nor the RAs have informed the participants of anything about the research. This is an obvious violation of the first principle because the students (participants) should have the right to know the relevant information about the usage of their own profile data and the purpose the research, so they can then decide to participate in the project or not. But in the case of the T3 project, they were not even aware of their profile data having been used.

2 Beneficence

Another important principle of conducting ethical research is to make sure the research will do no harm to all participants and at the same time maximize possible benefits. One benefit of directly extracting students’ profile data from Facebook without direct contact with students by T3 team is that the collected data would not

be biased or be biased to a lesser extent. In addition, the data's completeness and being longitudinal allows for close study on the change of the social network overtime.

While the large size of the data collected by the T3 team will doubtlessly benefit the further relevant research which can fully leverage the rich data, we cannot ignore the risk exposing the participants to potential trouble after the their profile data were released, because the re-identification was accomplished very soon after the release. Although the T3 team claimed that all the personal information they collected can be found on Facebook, but there are two possible added risks the students would be exposed after the re-identification. The first one is that there may be some personal information which was protected by the Facebook privacy setting now being made public. The second one is that the aggregate information from the whole cohort after the re-identification becoming public will be very likely to draw more attention and therefore expose the students to a higher lever risk.

3 Justice

The third principle of justice basically emphasizes that those who contribute to the research and those who benefit from the research should be distributed fairly. In the case of the T3 project, one issue related to the Justice principle is the compensation given to the participants. Students whose profile data were collected from their Facebook accounts were not compensated financially at all, and of course they did not participate in the project voluntarily because they were not even aware of their data having been used, so I am sure they would not feel having been compensated in other ways either.

The other issue about the Justice principle in this case is that the participants were comprised of only students from one college while the whole society may benefit from the research. So the group of participants are less representative, and their own profile data leaked out and were used for free, which eventually could contribute to some common benefit to the whole society. In this sense, the risks and benefits of the research seemed not to have been distributed fairly.

4 Respect for Law and Public Interest

The fourth principle of ethical research concerns about the interest of all stakeholders related to the research. The first implication of the principle is to obey relevant laws, contracts and terms, and the second implication of the principle is about the transparency and the accountability of the research.

For the T3 project, one possible issue related to the fourth principle may be the privacy violation. For example, the participants' personal information being used without their informed consent may be illegal. For one thing, the personal data were originally meant for social networking among friends on Facebook, but now they can be downloaded for various secondary uses without the authorization of the

participants. For another thing, the participants' housing information and email addresses were originally used by the university for its administration, but then they were given to the T3 team for the purpose of research without the authorization of the participants. Both incidents may be illegal because they clearly infringed on the participant's privacy.

Considering all the above violations of the four principles of ethical research, especially the privacy violation, I would be reluctant to use the dataset for my own research. If the dataset is indispensable to my current research, I will try my best to find out any possible way to contact one or more of the participants and thoroughly explain to them what the dataset are used for, and I will only use the data after I obtain their permissions.

References:

- [1] Lewis, K., Kaufman, J., Gonzalez, M., Wimmer, A., & Christakis, N. (2008). Tastes, ties, and time: A new social network dataset using Facebook.com. *Social networks*, 30(4), 330-342.
- [2] Wimmer, A., & Lewis, K. (2010). Beyond and below racial homophily: Erg models of a friendship network documented on facebook. *American Journal of Sociology*, 116(2), 583-642.
- [3] Lewis, K., Gonzalez, M., & Kaufman, J. (2012). Social selection and peer influence in an online social network. *Proceedings of the National Academy of Sciences*, 109(1), 68-72.
- [4] Zimmer, M. (2010). "But the data is already public": on the ethics of research in Facebook. *Ethics and information technology*, 12(4), 313-325.